

УДК 54.02+519.852.6

## Дескрипторы водородной связи в компьютерном молекулярном дизайне

О. А. Раевский

*ОЛЕГ АЛЕКСЕЕВИЧ РАЕВСКИЙ — доктор химических наук, профессор, заведующий Отделом компьютерного молекулярного дизайна, заместитель директора Института физиологически активных веществ РАН, председатель Российской секции Международного общества QSAR and Modelling Society. Область научных интересов: компьютерный молекулярный дизайн, физико-химические дескрипторы молекулярной структуры, методология взаимосвязи «структура—свойство (активность)».*

*142432 г.Черноголовка Московской обл., Северный проезд, 1, Институт физиологически активных веществ РАН, тел. (факс) (495)785-70-24, E-mail raevsky@ipac.ac.ru*

В современном компьютерном молекулярном дизайне эффективных химических соединений с полезными свойствами, основанном на концепции взаимосвязи молекулярной структуры и свойства (биологической активности), можно выделить три главных этапа работы.

1.Формирование обучающей выборки соединений с заданным свойством (активностью).

2.Описание молекулярной структуры исследуемых соединений.

3.Установление взаимосвязи «структура—свойство (биологическая активность)» с последующим созданием устойчивых прогностических математических моделей.

Вполне понятно, что качество работы по всем этим трем направлениям имеет важное значение для конечного суммарного результата компьютерного моделирования веществ с заданными полезными свойствами.

Количественное описание молекулярной структуры химических соединений в процессе компьютерного молекулярного дизайна осуществляется с помощью дескрипторов — независимых переменных, характеризующих особенности структуры молекул или их частей (заместители, фрагменты). Для корректного использования сведений о молекулярной структуре в компьютерном молекулярном дизайне необходимо иметь представление дескрипторов в форме, удобной для применения в вычислительном процессе и одновременно сохраняющей большую часть информации о собственно молекулярной структуре.

В литературе описаны и применяются в практике компьютерного молекулярного дизайна тысячи различных дескрипторов, отображающих самые тонкие детали молекулярной структуры. Одновременное использование всех этих дескрипторов бессмысленно ввиду того, что между ними существуют многочисленные корреляции. Отсюда ясна польза классификации дескрипторов молекулярной структуры.

Существуют различные способы классификации дескрипторов молекулярной структуры. Так, например, дескрипторы подразделяются на «чистые», описывающие какой-либо один эффект внутри- и межмолекулярных взаимодействий, и композитные, описывающие два и более эффекта. Другой тип классификации основывается на способах оценки дескрипторов (экспериментальные или теоретические). Систематическая классификация дескрипторов дана в публика-

циях [1,2]. Так, в [2] приводится схема нескольких уровней дескрипторов: дескрипторы элементного уровня, дескрипторы структурной формулы, дескрипторы электронной структуры, дескрипторы молекулярной формы и дескрипторы межмолекулярных взаимодействий. При таком представлении каждый последующий уровень включает информационное содержание предыдущего. Таким образом, получается, что дескрипторы межмолекулярных взаимодействий содержат всю информацию, заключенную во всех предыдущих уровнях описания молекулярной структуры. Фактически это означает, что концепция взаимосвязи «структура—свойство (активность)» является отражением взаимосвязи молекулярных взаимодействий химических соединений между собой и с биологическими объектами. Все предыдущие уровни описания молекулярной структуры (топологические дескрипторы, дескрипторы структурной формулы, квантовохимические дескрипторы) несут только часть информации о молекулярной структуре и могут быть эффективными только для описания достаточно простых свойств простых соединений (например, температуры кипения углеводородов).

Специфические межмолекулярные взаимодействия рассчитываются на основе предположения об аддитивном вкладе взаимодействий различных типов: стерических, электростатических контактов и образования водородной связи. Настоящая статья посвящена анализу дескрипторов водородной связи, используемых в компьютерном молекулярном дизайне.

Водородная связь имеет решающее значение в образовании многих молекулярных и ионных комплексов практически во всех биологических процессах, включая ферментативный катализ. Наиболее яркий пример, показывающий важнейшую роль водородной связи в биологических системах, представлен Полингом при описании процесса репликации ДНК [3]. Этот процесс сопровождается образованием комплексов пуриновых и пиримидиновых оснований путем формирования двух водородных связей — связывание аденина и тимина или трех — связывание гуанина и цитозина. Проведенный Полингом даже весьма приближенный расчет разности свободных энергий этих процессов комплексообразования (принималось, что каждая водородная связь дает выигрыш в энергии, равный 20 кДж/моль) продемонстрировал, что указанная система водородных связей практически исключает вероятность других комбинаций связывания.

Несмотря на столь важную роль водородной связи в химических и биологических процессах, учет этого межмолекулярного взаимодействия даже в настоящее время проводится преимущественно на примитивном уровне оценки индикаторных переменных, например числа атомов азота и кислорода с присоединенными атомами водорода (см. правило Липински [4]). Разумной альтернативой этому и другим качественным подходам может быть только описание образования водородной связи на основе термодинамических функций процесса комплексообразования:

$$\Delta G = -RT \ln K = \Delta H - T\Delta S \quad (1)$$

где  $\Delta G$  — разность свободных энергий;  $T$  — абсолютная температура;  $K$  — константа равновесия;  $\Delta H$  — разность энтальпий;  $\Delta S$  — разность энтропий.

Развитие эмпирических термодинамических шкал водородной связи, выводимых на основе экспериментальных значений термодинамических параметров, началось в 1960-е годы. При этом использовались два подхода: аддитивно-мультипликативный и мультипликативный. В случае аддитивно-мультипликативного подхода изменение энтальпии при образовании водородной связи вначале описывалось четырьмя параметрами:

$$-\Delta H = E_A E_B + C_A C_B \quad (2)$$

где индекс А относится к кислоте Льюиса, В — к основанию Льюиса.

Затем была предложена более заформализованная формула с шестью параметрами:

$$-\Delta H = E_A E_B + C_A C_B + T_A T_B \quad (3)$$

Каждый терм уравнения (3) соответствует конкретному типу межмолекулярных взаимодействий:  $E_A E_B$  — электростатическому,  $C_A C_B$  — ковалентному,  $T_A T_B$  — комплексообразованию с переносом заряда [5–11]. Использование указанного аддитивно-мультипликативного подхода было достаточно успешным в физической органической химии. Применение же его в молекулярном компьютерном дизайне для установления корреляции «структура—свойство (биологическая активность)» затруднено из-за отсутствия в подавляющем большинстве случаев точных сведений о молекулярной структуре одного из партнеров по молекулярному взаимодействию (например, рецепторов, конкретной биологической мишени).

Более широкое распространение получил метод описания термодинамических параметров водородной связи в рамках мультипликативного подхода. В этом случае энтальпия водородной связи описывается уравнением [12]:

$$\Delta H_{ij} = \Delta H_{11} P_i E_j \quad (4)$$

где  $\Delta H_{11}$  — изменение энтальпии при образовании водородной связи между фенолом и диэтиловым эфиром в четыреххлористом углероде (стандартная система);  $P_i$  — безразмерный фактор  $i$ -го донора водородной связи;  $E_j$  — фактор  $j$ -го акцептора водородной связи.

Для описания изменения энтальпии  $\Delta H_{ij}$  можно использовать также донорный ( $E_d$ ) и акцепторный ( $E_a$ )

энтальпийные факторы водородной связи, имеющие противоположные знаки [13,14]:

$$\Delta H = k_1 E_d E_a \quad (5)$$

Примеры применения этого подхода к оценке акцепторной способности водородной связи для обширного массива фосфорильных соединений представлены в публикациях [15–17].

Аналогичное уравнение было предложено и для описания изменения свободной энергии при образовании водородной связи [18]:

$$\Delta G = k_1 C_d C_a + k_0 \quad (6)$$

где донорный ( $C_d$ ) и акцепторный ( $C_a$ ) — свободноэнергетические факторы водородной связи.

Развитие и успешное применение рассматриваемого подхода к количественному описанию водородной связи возможно лишь на базе обширных экспериментальных данных по термодинамике водородной связи. Наличие массива таких данных позволило получить путем расчета статистически устойчивые термодинамические параметры водородной связи [19]. Из многих тысяч систем с водородной связью были отобраны 936 систем, которые удовлетворяли следующим требованиям: 1) стехиометрия комплексов с водородной связью 1:1; 2) растворитель для всех систем — четыреххлористый углерод; 3) для комплексов имелись данные и по  $\Delta H$ , и по  $\Delta G$ ; 4) значения как  $\Delta H$ , так и  $\Delta G$  установлены экспериментальными методами.

Был выработан специальный алгоритм расчета энтальпийных и свободноэнергетических параметров водородной связи с выбором гексаметилфосфотриамида в качестве стандартного акцептора (встречался 88 раз среди изученных 936 комплексов) и фенола в качестве стандартного донора водородной связи (встречался 119 раз). По алгоритму, описанному в работе [19], были рассчитаны донорные факторы водородной связи для 163 соединений и акцепторные факторы водородной связи для 195 соединений. Проверка полученных значений факторов водородной связи проводилась путем сопоставления экспериментальных значений энтальпий и свободных энергий с рассчитанными на основе мультипликативного подхода и дала следующие результаты (см. [19]):

$$\Delta H_{\text{рас}} = -0,27(\pm 0,45) + 1,00(\pm 0,02)\Delta H_{\text{экс}} \quad (7)$$

$$n = 936, r = 0,954, s = 2,70, F = 9553$$

$$\Delta G_{\text{рас}} = -0,07(\pm 0,12) + 1,00(\pm 0,01)\Delta G_{\text{экс}} \quad (8)$$

$$n = 936, r = 0,984, s = 1,11, F = 28556$$

Здесь и далее  $n$  — число используемых для нахождения корреляции соединений,  $r$  — коэффициент корреляции,  $s$  — стандартное отклонение,  $F$  — критерий Фишера.

Статистические параметры уравнений (7) и (8) демонстрируют, что оценки факторов водородной связи обеспечивают получение достоверных значений энтальпии и свободной энергии водородного связывания.

На рис.1 приведена шкала водородной связи, позволяющая количественно сопоставлять силу акцепторов водородной связи между собой и силу доноров водородной связи.

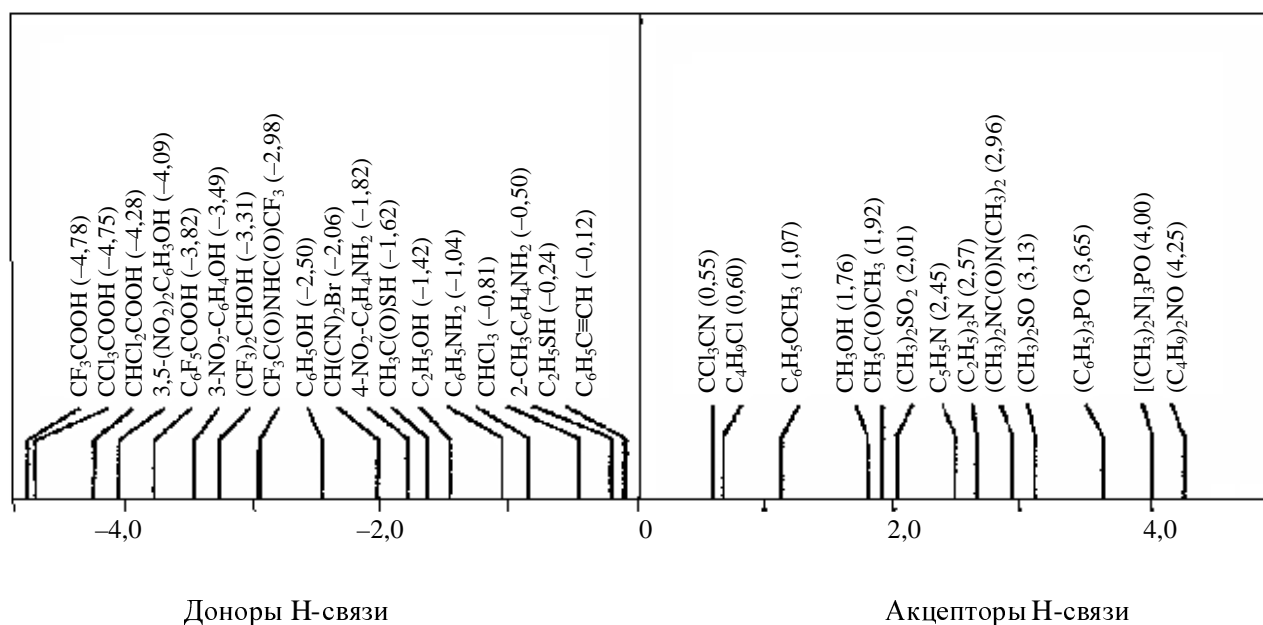


Рис. 1. Графическое сопоставление донорной и акцепторной способности водородной связи для простых органических соединений. В скобках приведены значения безразмерных свободноэнергетических факторов водородной связи

С использованием экспериментальных параметров термодинамики водородной связи была проведена аналогичная оценка свободноэнергетических факторов для 414 доноров и 1298 акцепторов водородной связи [20,21]. По значениям  $S_d$  для 414 доноров и  $S_a$  для 1298 акцепторов водородной связи была рассчитана свободная энергия для 537372 комплексов с водородной связью.

В работе [19] было установлено, что интервалы значений энтальпийных и свободноэнергетических факторов водородной связи для различных функциональных групп перекрываются (см. таблицу). Это свидетельствует о том, что заместители при донорном или акцепторном атомах в молекулах играют исключительно важную роль в формировании способности этих атомов к образованию водородной связи.

В литературе описаны также аналогичные методы расчета свободноэнергетических и энтальпийных параметров водородной связи на основе мультипликативного подхода [22–24]. Так, например, в работе [24] установлены факторы водородной связи для 150 доноров и 500 акцепторов, что дает возможность рассчи-

тать константы образования 75000 комплексов с водородной связью.

Интенсивные исследования по количественному описанию водородной связи ведутся в Институте физиологически активных веществ РАН (ИФАВ РАН). Систематическая работа Отдела компьютерного молекулярного дизайна Института (ОКМД ИФАВ РАН) с 1980 года привела к составлению обширной базы экспериментальных данных по термодинамике водородной связи и созданию программного комплекса НУВОТ (НУdrogen VOnd Thermodynamics) [20, 21, 25].

На рис. 2 представлено главное меню программного комплекса НУВОТ. Текущая версия НУВОТ содержит три программных модуля: базу термодинамических параметров водородной связи (около 16000 систем); базы энтальпийных и свободноэнергетических факторов водородной связи (более 25000) и программу расчета энтальпийных и свободноэнергетических факторов водородной связи. Программный комплекс НУВОТ функционирует в операционных системах Windows, UNIX, LINUX.

Таблица

Минимальные и максимальные значения факторов свободной энергии водородной связи для различных функциональных группировок

Доноры водородной связи	$S_d^{\min}$	$S_d^{\max}$	Акцепторы водородной связи	$S_a^{\min}$	$S_a^{\max}$
$\text{R}^1\text{R}^2\text{R}^3\text{C}-\text{H}$	-2,06	-0,12	$\text{R}-\text{O}-\text{R}$	0,78	2,01
$\text{RS}-\text{H}$	-1,92	-0,24	$\text{R}-\text{S}-\text{R}$	0,57	1,51
$\text{R}^1\text{R}^2\text{N}-\text{H}$	-3,52	-0,50	$\text{R}^1\text{R}^2\text{R}^3\text{N}$	0,58	3,65
$\text{PhO}-\text{H}$	-4,24	-0,76	$\text{RC}\equiv\text{N}$	0,55	2,77
$\text{AlkO}-\text{H}$	-3,50	-1,06	$\text{R}^1\text{R}^2\text{R}^3\text{P}=\text{O}$	2,40	4,02
$\text{RC(O)O}-\text{H}$	-4,78	-2,50	$\text{R}^1\text{R}^2\text{S}=\text{O}$	1,16	3,25

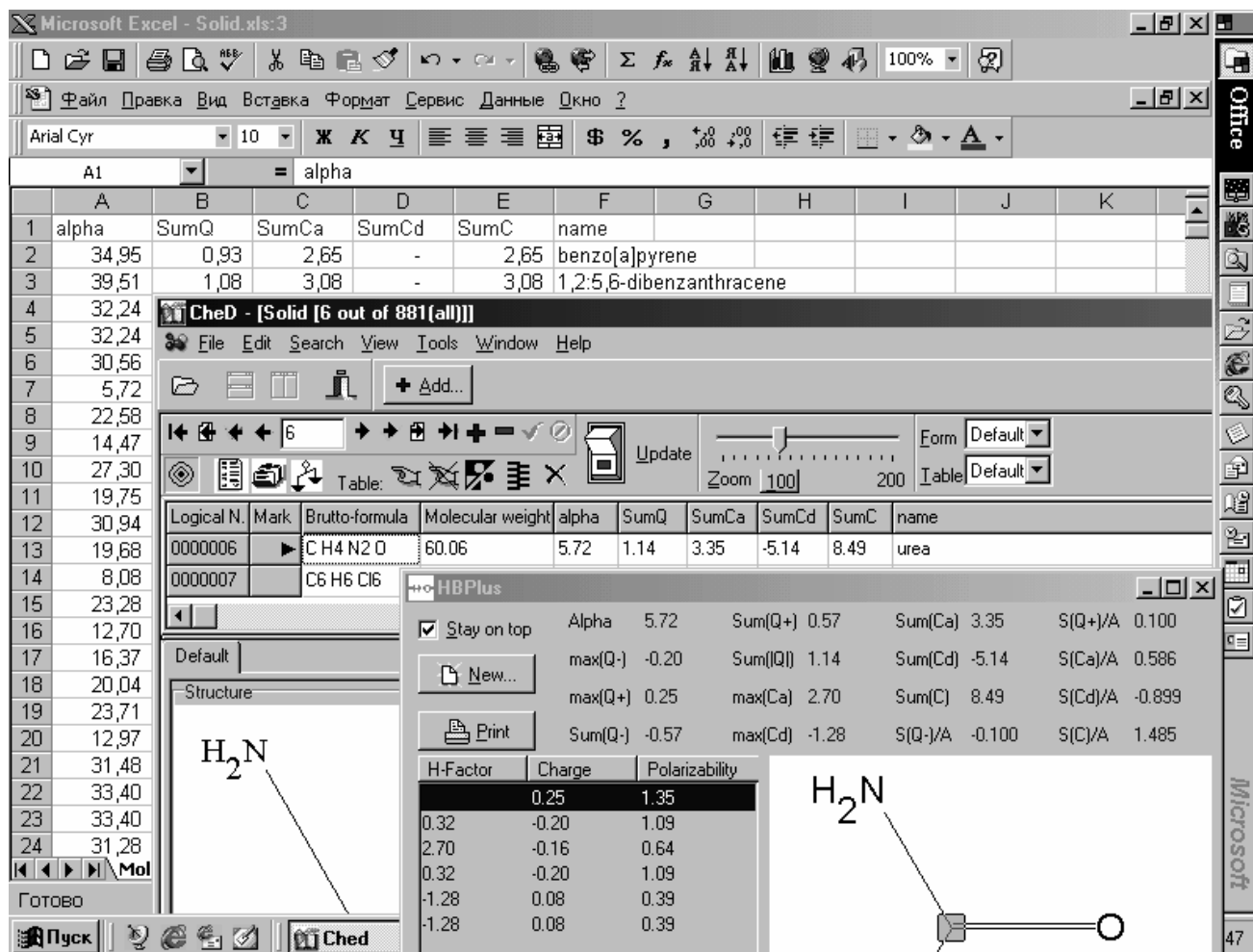


Рис. 2. Главное меню программного комплекса НУВОТ

В подходе QSAR (количественная связь «структура–активность») моделирования, развиваемом в ОКМД ИФАВ РАН, любое свойство  $P$  рассматривается как функция параметров трех межмолекулярных взаимодействий: стерического, электростатического и образования водородных связей [26]:

$$P = f(\alpha, \Sigma q, \Sigma C) \quad (9)$$

где  $\alpha$  — молекулярная поляризуемость, описывающая стерические взаимодействия;  $\Sigma q$  — сумма зарядов, описывающая электростатические взаимодействия;  $\Sigma C$  — сумма факторов водородной связи.

Ниже представлены примеры использования двухмерных дескрипторов водородной связи, рассчитываемых с помощью программного комплекса НУВОТ для установления количественной взаимосвязи «структура–свойство (биологическая активность)».

#### Модели липофильности органических соединений

Оценка распределения химических соединений и лекарств в системе октанол–вода — наиболее распространенный способ установления липофильности. Этот параметр является исключительно важным для суждения о транспортных возможностях веществ в биологических объектах и поэтому повсеместно используется в QSAR.

Работы по созданию физико-химических моделей липофильности, основанные на стерических и электростатических взаимодействиях и водородного связывания ведутся с начала 1980-х годов [27, 28]. В 1995 г. было предложено уравнение, в котором в качестве дескрипторов для описания коэффициентов распределения химических соединений в системе октанол–вода ( $\log P_{ow}$ ) используются молекулярный объем и факторы водородной связи [29]. Позднее молекулярный объем был заменен на молекулярную поляризуемость ( $\alpha$ ). В результате была получена великолепная корреляция  $\log P_{ow}$ , установленная на базе 2850 простых монофункциональных органических соединений, с применением только двух дескрипторов — поляризуемости и акцепторного фактора водородной связи:

$$\log P_{ow} = 0,267\alpha - 1,00\Sigma C_a \quad (10)$$

$$n = 2850, r = 0,970, s = 0,23$$

Это уравнение было использовано для расчета липофильности с помощью программы SLIPPER-98 [30]. Следует отметить две важные особенности этого уравнения: 1) несмотря на то, что значительная часть соединений обучающей выборки содержала водородно-связанные донорные группировки, вклад этого фактора в липофильность не был установлен; 2) уравнение не имеет свободного члена, что свидетельствует

об адекватной оценке липофильности на основе указанных двух дескрипторов.

Анализ уравнения (10) приводит к следующим заключениям. Ситуация, когда оба термина уравнения имеют одинаковые величины, означает, что химическое соединение будет одинаково распределяться в двух фазах (октанол и вода). Далее, поскольку поляризуемость имеет только положительное значение, и следовательно, этот параметр вносит положительный вклад в липофильность, то увеличение поляризуемости ведет к возрастанию концентрации вещества в октанольной фазе. Акцепторные факторы водородной связи также могут быть только положительными, но поскольку при этом факторе в уравнении (10) стоит отрицательный знак, то значит увеличение акцепторной способности водородной связи будет приводить к возрастанию концентрации вещества в водной фазе.

Применение в программе SLIPPER-98 уравнения (10) продемонстрировало хорошую предсказательную его силу в отношении липофильности, однако в случае сложных химических соединений, в частности веществ-лекарств, содержащих в составе молекул несколько функциональных групп, возможны существенные отклонения от экспериментальных значений. Для повышения точности расчетов в формулу (10) помимо физико-химических дескрипторов было также включено экспериментальное значение липофильности для ближайшего структурно родственного соединения [31]:

$$\log P_{ow(i)} = \log P_{ow(mm)} + 0,267 (\alpha_i - \alpha_{mm}) - 1,00(\Sigma C_{a(i)} - \Sigma C_{a(mm)}) \quad (11)$$

где величины с индексом *mm* относятся к родственному соединению.

Проверка эффективности расчетов липофильности по формуле (11) с помощью программного комплекса SLIPPER-200 на массиве, содержащем 10937 органических веществ, в том числе лекарств [32], показала хорошее согласие между экспериментальными и рассчитанными значениями, коэффициент корреляции составил 0,972.

#### Расчет растворимости химических соединений в воде

Растворимость в воде является очень важным свойством для лекарственных соединений — от него зависит усваиваемость лекарства, проницаемость через биологические мембраны, распределение в различных системах воздух—вода—растворитель.

Еще в 1968 году на основе данных для 150 жидких соединений Ганч, пионер QSAR, логически предположил, что растворимость (*S*) и липофильность связаны обратно пропорциональной зависимостью [33]. С.Ялковский с сотр. развил эту идею и предложил следующее количественное соотношение [34]:

$$\log S = -1,07 \log P_{ow} + 0,67 \quad (12)$$

Очевидно, что если  $\log P_{ow}$  (липофильность) количественно описывается с помощью молекулярной поляризуемости и факторов водородной связи, то эти дескрипторы должны быть эффективными и для описания растворимости в воде. Это было подтверждено кор-

реляционным анализом обучающей выборки по растворимости, содержащей 630 жидких веществ [35]:

$$\log S = 0,58 - 0,305\alpha + 1,16\Sigma C_a \quad (13)$$

$$n = 630, r = 0,936, s = 0,59$$

С учетом того результата, выявленного при изучении 45 нейтральных полярных жидких соединений, что кроме поляризуемости и акцепторной способности водородной связи на растворимость в воде может оказать существенное влияние донорная способность водородной связи ( $\Sigma C_d$ ) [21,36], получена следующая корреляция растворимости жидких веществ в воде:

$$\log S = 0,43 - 0,298\alpha + 1,09\Sigma C_a - 0,30\Sigma C_d \quad (14)$$

$$n = 630, r = 0,947, s = 0,54$$

В случае кристаллических химических соединений и лекарств значительный вклад в растворимость может вносить энергия кристаллической решетки. Для учета этого эффекта С.Ялковский предложил использовать в качестве дополнительного дескриптора температуру плавления ( $T_{пл}$ ) [37]:

$$\log S = 0,38 - 0,0096(T_{пл} - 25) - 1,022 \log P_{ow} \quad (15)$$

$$n = 1026, r = 0,980, AAE = 0,38, RMSE = 0,53$$

$$\text{где } AAE = (\Sigma |\log S_{расч} - \log S_{экс}|) / n;$$

$$RMSE = \sqrt{[\Sigma (\log S_{расч} - \log S_{экс})^2] / n}.$$

Следует, однако, отметить, что изученный массив соединений содержал 497 жидких веществ, т.е. составлял почти половину соединений, включенных в обучающую выборку, следовательно, в целом обучающая выборка не отражала достаточно полно вклад кристаллической фазы. В работе [38] статистическая оценка данных по растворимости массива из 1063 соединений, содержащего исключительно кристаллические соединения и лекарства, методом прямого регрессионного анализа с учетом липофильности и температуры плавления привела к выводу следующего соотношения:

$$\log S = 0,03 - 0,0091(T_{пл} - 25) - 0,968 \log P_{ow} \quad (16)$$

$$n = 1063, r = 0,923, AAE = 0,60, RMSE = 0,79$$

Как видно, статистические параметры уравнения (16) по сравнению с параметрами уравнения (15) значительно хуже.

Альтернативным методом расчета растворимости кристаллических соединений в воде является подход, основанный на концепции сходства и использования физико-химических дескрипторов, находимых программой NYBOT. В этом случае используются коэффициенты из уравнения, связывающего растворимость жидких соединений с физико-химическими параметрами, а экспериментальное значение растворимости в воде ближайшего структурно родственного соединения учитывает эффект кристаллической решетки [38,39]:

$$\log S_i = \log S_{nn} - 0,275(\alpha_i - \alpha_{nn}) + 0,895(\Sigma C_{a,i} - \Sigma C_{a,nn}) + 0,325(\Sigma C_{d,i} - \Sigma C_{d,nn}) \quad (17)$$

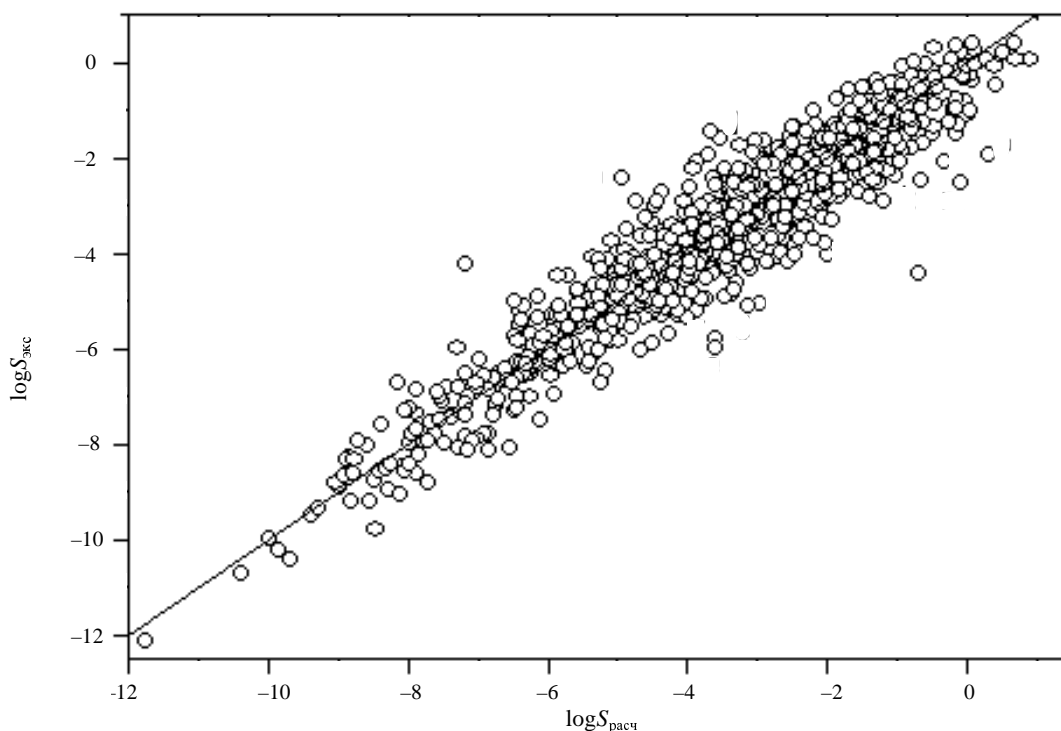


Рис.3. Сопоставление экспериментальных и рассчитанных по уравнению (17) показателей растворимости в воде для 1063 химических соединений

С использованием дескрипторов для ближайших структурно родственных соединений, имеющих близкие значения физико-химических параметров, удалось получить хорошую корреляцию для тех же 1063 кристаллических химических соединений и лекарств, коэффициент корреляции  $r = 0,951$ , ААЕ = 0,47, RMSE = 0,64. Графически эта корреляция представлена на рис.3.

*Расчет проницаемости через биологические мембраны*

А. Базальная проницаемость в красные кровяные клетки человека

Напомним, что базальная проницаемость (basal permeability, BP) — это перенос вещества в отсутствие какого-либо специфического транспортного пути (осуществляется в случае ингибирования механизмов активного транспорта веществ). Математическое моделирование этого процесса на базе экспериментальных данных по проницаемости 10 простых химических соединений с использованием традиционного дескриптора  $\log P_{ow}$  дало весьма скромные результаты [40]:

$$\log P_{BP} = -3,31 + 0,79 \log P_{ow} \quad (18)$$

$$n = 10, r = 0,619, s = 1,86$$

Введение же только фактора донорной способности водородной связи привело к великолепной корреляции:

$$\log P_{BP} = -0,70 + 1,08 \Sigma C_d \quad (19)$$

$$n = 10, r = 0,983, s = 0,43$$

Б. Проницаемость веществ через клетки *alga Chara ceratophylla*

Для данного вида проницаемости ( $P$ ) веществ была получена хорошая корреляция также на основе факторов водородной связи [40].

$$\log P = -0,83 + 0,59 \Sigma C_d \quad (20)$$

$$n = 27, r = 0,903, s = 0,49$$

В. Проницаемость через кожу человека

В работе [40] было найдено также корреляционное соотношение для проницаемости веществ через кожу  $\log K_p$ :

$$\log K_p = -4,88 - 0,31 \Sigma C_a + 0,23 \Sigma C_d \quad (21)$$

$$n = 22, r = 0,912, s = 0,50$$

(Отметим, что *Chara ceratophylla* — вид водоросли.)

Г. Распределение лекарств в липосомах

Подробное QSAR исследование распределения 54 лекарств в липосомах представлено в работе [41]. Получено регрессионное уравнение, основанное на НУВОТ-дескрипторах, применительно к лекарствам-веществам нейтрального ( $\log P_{ui}$ ), кислотного ( $\log P_{(-)}$ ) и основного ( $\log P_{(+)}$ ) характера:

$$\log P_{ui} = 3,14 + 0,080\alpha - 0,22 \Sigma C_a - 0,91 \Sigma q \quad (22)$$

$$\log P_{(-)} = 1,76 + 0,059\alpha - 1,17 \Sigma C_a + 0,84 \Sigma q_{(-)} \quad (23)$$

$$\log P_{(+)} = 2,43 + 0,075\alpha - 0,08 \Sigma C_a - 0,93 \Sigma q_{(+)} \quad (24)$$

$$n = 54, r = 0,925, s = 0,37$$

где  $\alpha$  — молекулярная поляризуемость;  $\Sigma q$  — суммарный заряд.

### Расчет адсорбции лекарств в организме человека

В рамках работ по созданию лекарств в последнее десятилетие были выполнены систематические исследования, позволившие выявить вклад физико-химических свойств веществ в адсорбцию их в организме. Из физико-химических параметров соединений Липински [4] выделил четыре фактора, имеющих решающее значение для адсорбции лекарства, — это молекулярная масса, липофильность ( $\log P_{ow}$ ), число водородно-связанных доноров и число водородно-связанных акцепторов — и сформулировал известное под названием «правило пяти». Согласно этим правилам, низкая адсорбция или слабая проницаемость веществ реализуется, когда имеется более пяти водородно-связанных доноров (группы OH и NH), относительная молекулярная масса более 500,  $\log P > 5$  и число водородно-связанных акцепторов (количество атомов азота и кислорода в молекуле) более 10. Следует отметить, что дескрипторы, которые предложил использовать Липински, описывают взаимодействия атомов очень грубо. Так, стерические взаимодействия гораздо лучше описываются молекулярной поляризуемостью, а не молекулярной массой, а число водородно-связанных атомов слишком приблизительно характеризует способность молекул к образованию водородной связи.

На рис.4 представлена зависимость между числом водородно-связанных акцепторов в молекуле  $N_a$  и суммой термодинамических акцепторных факторов водородной связи  $\Sigma C_a$ . Из рисунка видно, что при одном и том же числе водородно-связанных акцепторных атомов энергия водородной связи меняется в очень широком интервале, так что дескрипторы, использованные Липински, дают возможность выявить только тенденцию влияния физико-химических дескрипторов на проницаемость и адсорбцию веществ [42].

В исследовании [43] подробно обсужден материал по проницаемости Сасо-2 (энтероцитоподобные клетки). Здесь были использованы различные стерические параметры (молекулярная масса, молекулярный объем вещества, полярные и неполярные части поверхности молекулы), число водородно-связанных доноров, водородно-связанные донорные и акцепторные факторы и некоторые физико-химические свойства. Было установ-

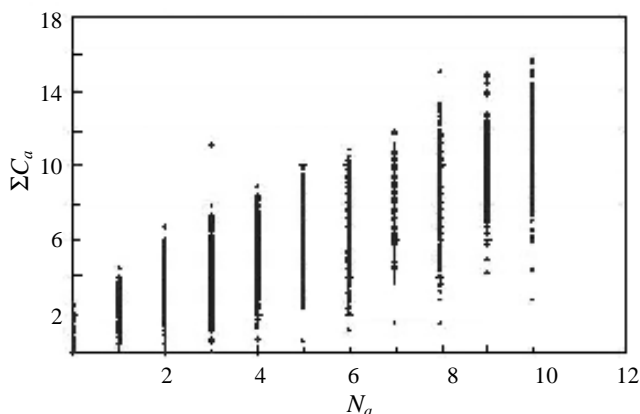


Рис.4. Зависимость между числом акцепторов  $N_a$  в молекуле и суммой свободноэнергетических факторов  $\Sigma C_a$  водородной связи

лено, что комбинация стерических дескрипторов молекулярной структуры и дескрипторов водородной связи обеспечивает разумную корреляцию проницаемости вещества.

Наконец, в работе [44] были использованы НУВОТ-дескрипторы для количественной оценки оральной адсорбции веществ в организме человека 31 лекарства с пассивным механизмом транспорта. Нелинейная модель, основанная только на одном дескрипторе, характеризующем суммарно донорную и акцепторную способности к образованию водородных связей ( $\Sigma C_{ad}$ ), весьма корректно описывает адсорбцию соединения:

$$A = 1/(1 + 10^{-[5,02 - 0,31\Sigma C_{ad}]}) \quad (25)$$

$$n = 31, r = 0,943, s = 0,12$$

Это уравнение показывает, что лекарственные соединения с  $\Sigma C_{ad} < 10$  полностью адсорбируются в организме человека, а с  $\Sigma C_{ad} > 22$  адсорбируются чрезвычайно слабо. Разделение композитного дескриптора  $\Sigma C_{ad}$  на «чистые» донорные и акцепторные факторы водородной связи позволяет достигнуть очень хороших статистических параметров:

$$A = 1/(1 + 10^{-[5,05 - 0,36\Sigma C_a + 0,26\Sigma C_d]}) \quad (26)$$

$$n = 31, r = 0,974, s = 0,09$$

Позднее аналогичное исследование было выполнено с использованием 100 лекарств, адсорбция которых осуществляется не только с помощью пассивного транспорта, но и за счет других специфических взаимодействий [45].

Яркий пример конкретного использования в ОКМД ИФАВ РАН двухмерных факторов водородных связей в компьютерном молекулярном дизайне для оптимизации свойств лекарств продемонстрирован на рис.5, где представлены рассчитанные ( $\log P_{ow(расч)}$  и  $\log S_{w(расч)}$ ) и экспериментальные ( $\log P$  и  $\log S$ ) значения.

### Трёхмерные факторы водородной связи

Дескрипторы водородной связи, о которых шла речь в приведенном выше материале, основываются на термодинамических параметрах образования комплексов с одной водородной связью. Сложные органические молекулы, в том числе лекарственных соединений, обычно содержат несколько функциональных групп и образуют сложные комплексы с участием множества водородных связей. Отсюда очевидно, что для моделирования взаимодействия таких соединений и лекарств с биологическими мишенями необходимо использование трёхмерных дескрипторов водородной связи. Современные процедуры трёхмерного компьютерного дизайна включают расчеты силового поля для учета водородного связывания. Эти расчеты выполняются во многих областях исследований, включая конформационный анализ, идентификацию фармакофоров, докинг лигандов, дизайн лиганда *de novo*, сравнительный анализ молекулярных полей, идентификацию молекулярных участков для благоприятного связывания и т.д.

Попытки улучшить расчет вклада водородной связи в энергию взаимодействия лиганда с макромолекулой были предприняты Гудфордом в рамках метода, на-

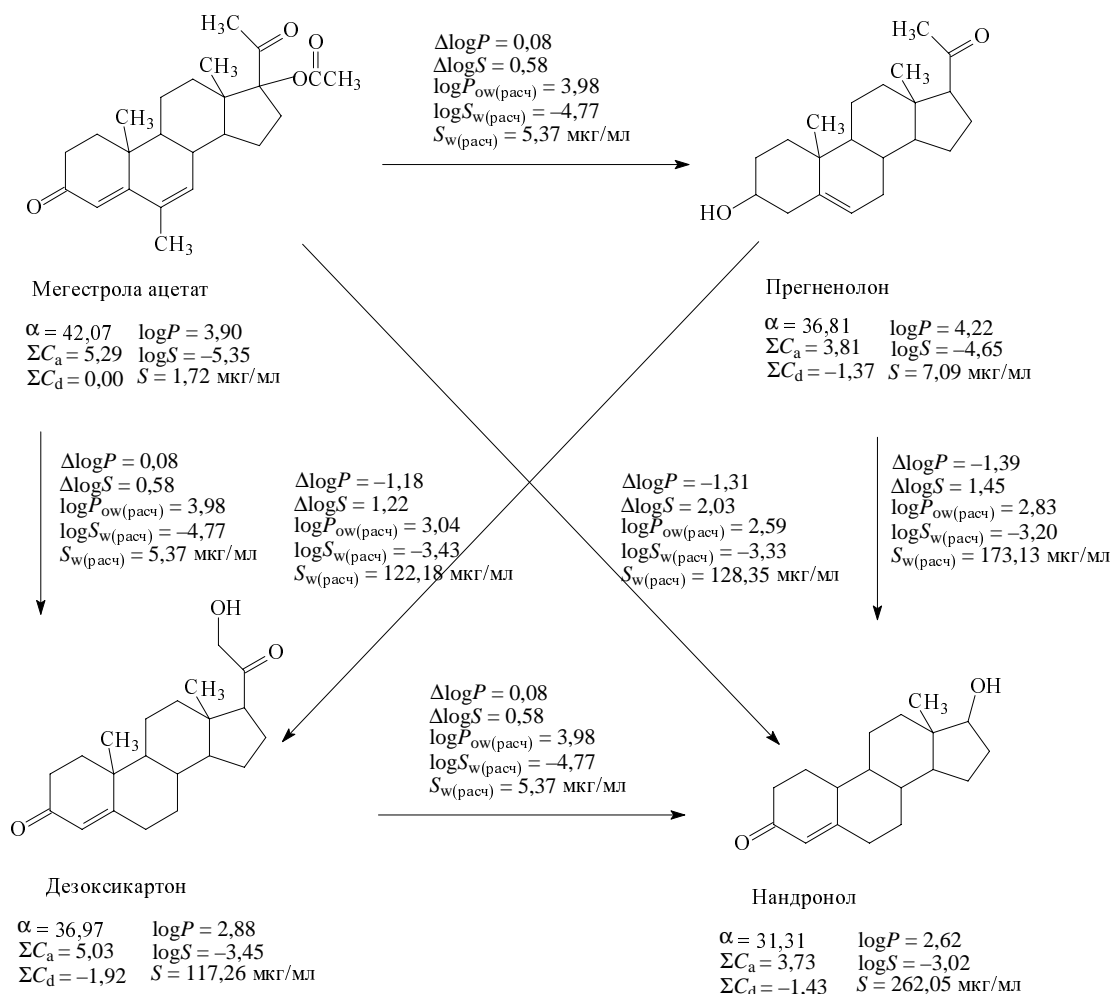


Рис.5. Оптимизация свойств лекарств на основе физико-химических дескрипторов НУВОТ, включающих свободноэнергетические водородно-связанные факторы

званного GRID [46—50]. Метод GRID параметризован на основе экспериментальных данных рентгеноструктурного анализа для небольших молекул кристаллических веществ и белков и реализуется в виде расчета взаимодействий выбранной пробной частицы (небольшой молекулы, такой как вода или аммиак) с макромолекулярной системой. Функция «8—6» адаптирована для описания этих взаимодействий:

$$E_r = C/r^8 - D/r^6 \quad (27)$$

где  $C = -3E_m r_m^8$ , ккал·Å<sup>8</sup>/моль;  $D = -4E_m r_m^6$ , ккал·Å<sup>6</sup>/моль;  $r$  — расстояние между акцепторным атомом и «тяжелым» донорным атомом;  $E_m$  — оптимальная энтальпия водородной связи в ккал/моль;  $r_m$  — оптимальная длина водородной связи.

В некоторых работах установление зависимости длины водородной связи от энергии основывается на потенциале «12—10» или «6—4».

Необходимо также упомянуть работу [51], где предложен расчет потенциалов водородной связи (МНВП), который основан на идее объединения двух-

мерного и трехмерного описания водородного связывания:

$$\text{МНВП} = \sum_{i=1}^N f_i f_{ct} (D_{ik}) f(U) \quad (28)$$

где  $k$  указывает точку (атом) в пространстве  $i$ -го рассматриваемого молекулярного фрагмента;  $N$  — полное число фрагментов, содержащих донорные и акцепторные центры водородной связи в молекуле;  $f_i$  —  $\alpha$  или  $\beta$  — свободноэнергетические сольватохромные параметры водородной связи  $i$ -го атома;  $f_{ct}$  — функция расстояния;  $D_{ik}$  — расстояние между фрагментом  $i$  и точкой  $k$ ;  $f(U)$  — угловая функция.

Следует оговорить, что применение этого подхода ограничено использованием сольватохромных параметров водородной связи. Кроме того, берутся фиксированные значения параметра функции расстояния, отвечающие оптимальным потенциалам водородной связи. Соответственно, в методе GRID фактически зафиксированы три потенциала водородной связи:  $-4,00$  ккал/моль в качестве оптимальной энергии водородной связи для  $O-H \cdots O$ ,  $-2,8$  ккал/моль в

качестве оптимальной энергии для O—H...N и -2 ккал/моль в качестве оптимальной энергии для N—H...N [47]. Эти значения потенциалов находятся в явном противоречии с данными по термодинамике водородной связи, собранными и систематизированными в программном комплексе НУВОТ. Данные НУВОТ свидетельствуют о том, что энтальпия водородной связи для одного и того же типа меняется в очень широких интервалах и существенно зависит от природы заместителя. Так, для O—H...O энтальпия водородной связи меняется в интервале 0,9—15,9 ккал/моль, для O—H...N — в интервале 0,5—10,5 ккал/моль, для N—H...N — в пределах 0,5—11,5 ккал/моль.

Для преодоления указанных недостатков метода GRID была проведена работа по установлению новых трехмерных потенциалов водородной связи [52]. С этой целью по данным рентгеноструктурного анализа были отобраны 58 «идеальных» комплексов с водородной связью, в которых углы между донорным «тяжелым» атомом, атомом водорода и атомами акцептора приходились на интервал 173—187°, а углы между ядрами акцепторных атомов, неподеленными парами электронов и атомом водорода были в интервале 170—190°. Оказалось 13 таких комплексов с водородной связью типа O—H...O, 19 комплексов O—H...N и 26 комплексов N—H...N.

Вследствие существенной разницы в ковалентных и ван-дер-ваальсовых радиусах атомов кислорода и азота соотношения между оптимальной энергией водородной связи  $E_m$  и оптимальной ее длиной  $r_m$  отскакивались отдельно для каждой из описанных выше трех систем. Ожидалось, что такие соотношения не будут линейными. В работе [52] была использована сигмоидная функция:

$$E_m = k_4 / (1 + 10^{k_5 + k_6 r_m}) \quad (29)$$

где  $k$  — коэффициенты, рассчитываемые из соответствующей системы уравнений.

Значения  $k_4$  в уравнении (29) могут быть зафиксированы и ограничены максимальными значениями энтальпии водородной связи. Перегруппировав уравнение (29) и упростив константы, авторы получили следующее соотношение:

$$r_m = k_7 \log[(k_4 - E_m)/(E_m)] + k_8 \quad (30)$$

Теперь стало возможным рассчитывать  $r_m$ , исходя из значений  $E_m$ .

По уравнениям (29) и (30) можно вычислить потенциалы водородной связи для каждой специфической пары атомов в молекуле. Примеры этих потенциалов для конкретных комплексов малых молекул приведены на рис.6.

В целом схема расчета энергии водородного связывания атомов в трехмерном пространстве в рамках указанного подхода может быть представлена следующим образом:

- 1) использование энтальпийных акцепторных и донорных факторов водородной связи, рассчитываемых программой НУВОТ на основе уравнения (5),
- 2) установление оптимального расстояния, соответствующего рассчитанной оптимальной энергии между атомами, образующими водородную связь, с помощью уравнения (30),
- 3) использование уравнения (27) для учета отклонений от идеального расположения атомов, включенных в образование водородной связи.

Пример результатов расчетов в рамках подхода, изложенного выше, приведен на рис.7. Здесь представлены взаимодействия пар оснований в А-форме двойной спирали РНК. Сопоставление полученных расчетных данных с результатами работы [47] показало, что расчеты на основе программного комплекса НУВОТ гораздо более корректно описывают эти сложные межмолекулярные взаимодействия.

Расчет трехмерных потенциалов водородной связи реализован в программе 3D НУВОТ [53].

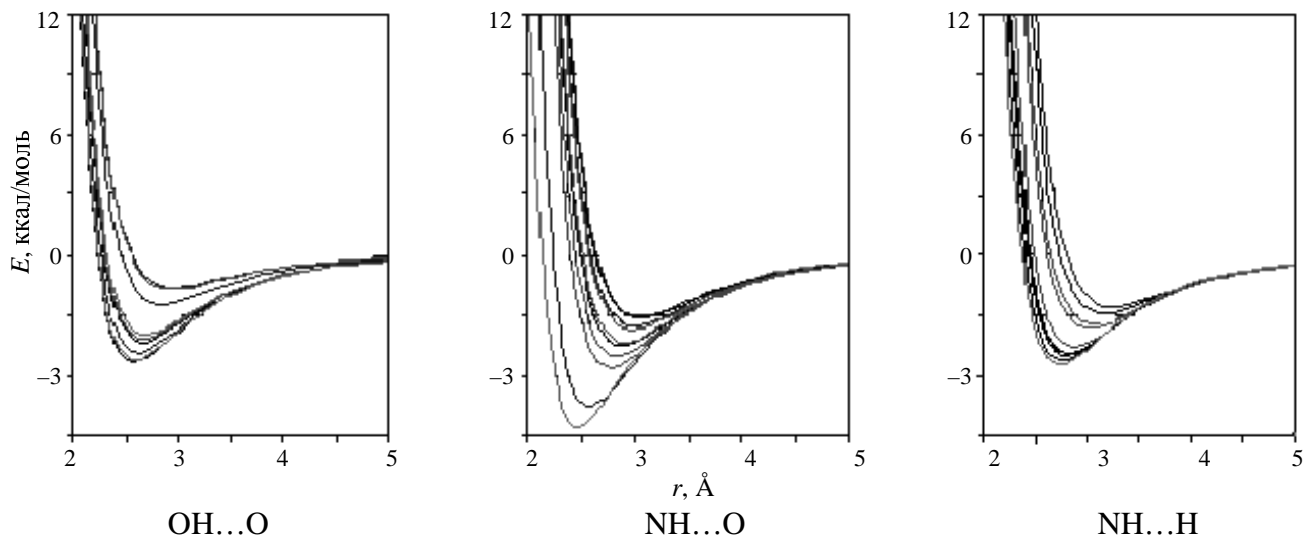


Рис.6. Потенциалы водородной связи для систем OH...O, NH...O и NH...N [52]

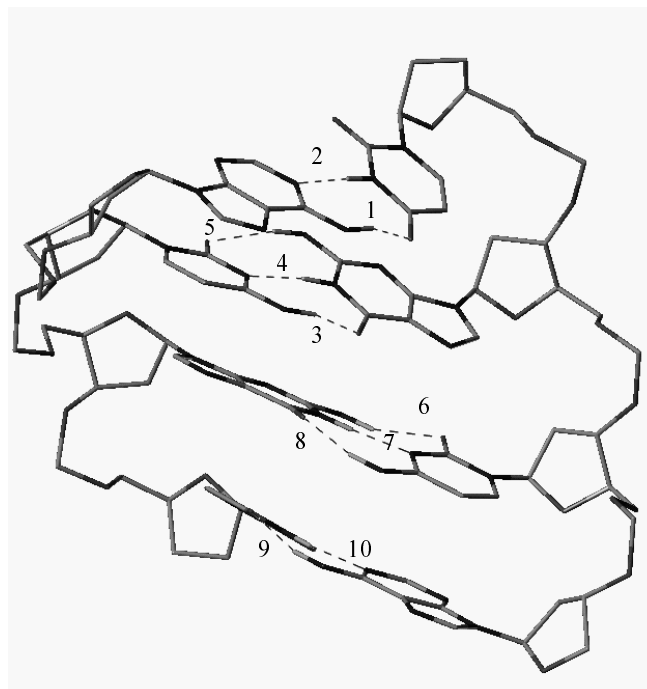
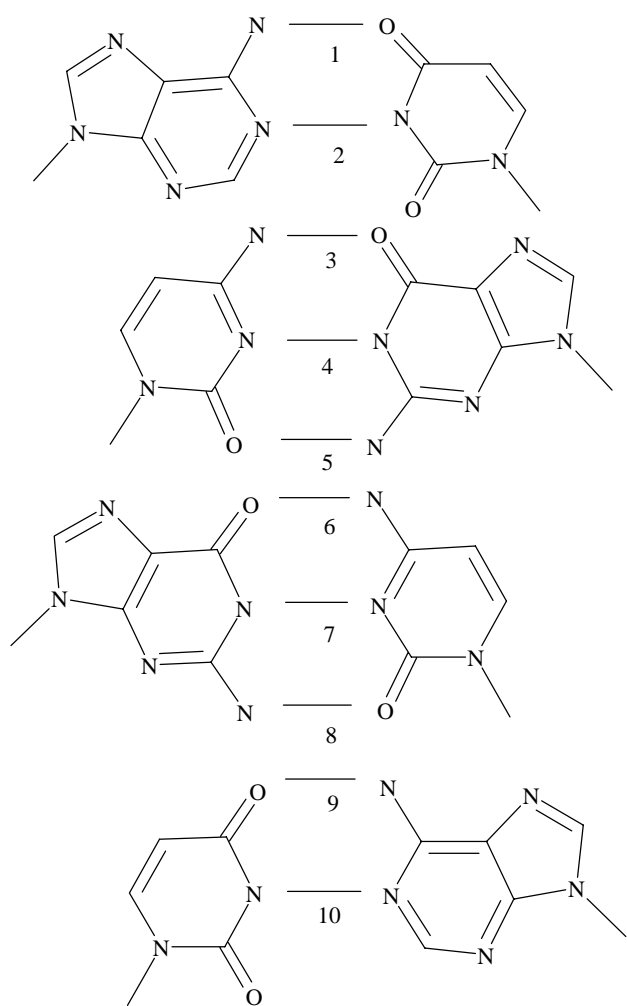


Рис.7. Набор водородных связей, реализуемых в А-форме двойной спирали РНК

Указанные оригинальные трехмерные потенциалы водородной связи легли в основу десяти новых трехмерных интегральных дескрипторов водородной связи [54]. В этой же работе они были успешно использованы для количественного описания пассивной диффузии 154 лекарств в организме человека.

Таким образом, к настоящему времени разработаны подходы к количественному описанию водородной связи. Эта задача возникает при моделировании взаимодействий сложных органических систем между собой и с биологическими объектами посредством водородного связывания — особого вида комплексобразования, реализуемого во многих химических и биохимических процессах. Двухмерные и трехмерные термодинамические дескрипторы водородной связи создают условия для эффективного учета вклада водородной связи в различные свойства химических соединений, в том числе в биологическую активность.

\* \* \*

Работа выполнена при поддержке Международного научно-технического центра (проект №888) и государственного контракта № 02.434.11.1014 с Федеральным

агентством по науке и инновациям Министерства образования и науки Российской Федерации.

#### ЛИТЕРАТУРА

1. Testa B., Kier L.B. *Med. Res. Rev.*, 1991, v. 11, p. 35.
2. Раевский О.А. *Успехи химии*, 1999, т. 68, вып. 6, с. 555.
3. Поллинг Л., Поллинг П. *Химия*. М.: Мир, 1978, с. 458.
4. Lipinski C.A., Lombardo F., Domini B.W., Feeney P.J. *Adv. Drug Del. Rev.*, 1997, v.23, p. 3.
5. Drago R.S., Wayland B.B. *J. Am. Chem. Soc.*, 1965, v. 87, p. 3571.
6. Drago R.S. *Structure and Bonding*. Springer-Verlag, Heidelberg, 1973, 227 p.
7. Drago R.S., Vogel G.C., Needham T.E. *J. Am. Chem. Soc.*, 1971, v. 93, p. 6014.
8. Guidry R.H., Drago R.S. *Ibid.*, 1973, v. 95, p. 759.
9. McMillan D.R., Drago R.S. *Inorg. Chem.*, 1972, v. 11, p. 872.
10. Drago R.S., Parr L.B., Chamberlain C.S. *J. Am. Chem. Soc.*, 1977, v. 99, p. 3203.
11. Kroeger M.K., Drago R.S. *Ibid.*, 1981, v. 103, p. 3250.
12. Йогансен А.В. *Теор. и эксп. химия*, 1971, т. 7, с. 302.
13. Раевский О.А., Авидон В.В., Новиков В.П. *Хим.-фарм. ж.*, 1982, т. 16, с. 968.

14. Раевский О.А., Новиков В.П. Там же, 1982, т. 16, с. 583.
15. Раевский О.А., Гилязов М.М., Лелин Я.А. Ж. общей химии, 1983, т. 53, с.563.
16. Раевский О.А., Гилязов М.М., Лелин Я.А. Там же, 1983, т. 53, с. 1720.
17. Раевский О.А., Гилязов М.М., Лелин Я.А. Там же, 1983, т. 53, с. 1724.
18. Раевский О.А., Григорьев В.Ю., Соловьев В.П. Хим.-фарм. ж., 1989, т. 23, с. 1294.
19. Raevsky O.A., Grigoriev V.Yu., Kireev D.B., Zefirov N.S. Quant. Str.—Act. Relat., 1992, v. 11, p. 49.
20. Raevsky O.A. J. Phys. Org. Chem., 1997, v. 10, p. 405.
21. Raevsky O.A. In: Computer-Aided Lead Finding and Optimization. Current Tools for Medicinal Chemistry. Eds. H.Waterbeemd, B.Testa, G.Folkers. Weinheim: VCH, 1997, p. 367.
22. Abraham M.H., Ibrahim A., Zissimos A.M., Zhao Y.H., Komer J., Reynolds D.P. Drug Discov. Today, 2002, v. 7, p. 1056.
23. Ruelle P. Perspect. Drug Discovery Des., 1999, v. 17, p. 61.
24. Abraham M.H. In Quantitative Treatments of Solute/Solvent Interactions. Theoretical and Computational Chemistry. V. 1. Eds. P.Polizer, J.S.Murrey. Amsterdam: Elsevier, 1994, p. 83.
25. <http://molpro.ipac.ac.ru>
26. Raevsky O.A. Mini Reviews in Medicinal Chemistry, 2004, v. 14, p. 1041.
27. Dearden J.C. Environ. Health Perspect., 1985, v. 61, p. 203.
28. van der Waterbeemd H., Testa D. Adv. Drug Res., 1987, v. 16, p. 85.
29. Raevsky O.A., Schaper K.-J., Seydel J. Quant. Struct.—Act. Relat., 1995, v. 14, p. 433.
30. Raevsky O.A., Trepalina E.P., Trepalin, S.V. SLIPPER — A New Program for Water Solubility, Lipophilicity and Permeability Prediction, Hydrogen bond contributions to properties and activities of chemicals and drugs. In: Molecular Modelling and Prediction of Bioactivity. Eds. K.Gundertofte, F.Jorgensen. Kluwer Academic/Plenum Publ., 2000, p. 489.
31. Raevsky O.A. SAR & QSAR in Environmental Research, 2001, v. 12, p. 367.
32. Raevsky O.A., Trepalin S.V., Trepalina E.P., Gerasimenko V.A., Raevskaja O.E., J. Chem. Inf. Comput. Sci., 2002, v. 42, p. 540.
33. Ganch C.F., Quinlan J.E., Lawrence G.L. J. Org. Chem., 1968, v. 33, p. 347.
34. Yalkovsky S.H., Morozovitch W., In: Medicinal Chemistry. Ed. E.J.Ariens. Academic Press, 1980, v. 9, p. 122.
35. Schaper K.-J., Kunz B., Raevsky O.A. QSAR & Comb. Sci., 2003, v. 22, p. 943.
36. Raevsky O.A., Schaper K.-J., van de Waterbeemd H., McFarland J.W. Hydrogen Bond Contributions to Properties and Activities of Chemicals and Drugs. In: Molecular Modelling and Prediction of Bioactivity. Eds. K.Gundertofte, F.Jorgensen. Kluwer Academic/Plenum Publ., 2000, p. 221.
37. Ran Y., He Y., Yang G., Jonson J.L.H., Yalkowsky S.H. Chemosphere, 2002, v. 48, p.487.
38. Raevsky O.A., Raevskaja O.E., Schaper K.-J. QSAR & Comb. Sci., 2004, v. 23, p.327.
39. Raevsky O., Andreeva E., Raevskaja O., Skvortsov V., Schaper K. SAR & QSAR in Environmental Research, 2005, v. 12, p. 191.
40. Raevsky O.A., Schaper K.-J. Europ. J. Med. Chem., 1998, № 33, p. 799.
41. Schaper K.-J., Zhang H., Raevsky O.A. Quant. Struct.—Act. Relat., 2001, v. 20, p. 46.
42. Waterbeemd H., Smith D.A., Beaumont K., Walker D.K. J. Med. Chem., 2001, v. 44, p. 1313.
43. Waterbeemd H., Gamenisch G., Folkers G., Raevsky O.A. Quant. Struct.—Act. Relat, 1996, v. 15, p. 480.
44. Raevsky O.A., Fetisov V.I., Trepalina E.P., McFarland J.W., Schaper K.-J. Ibid., 2000, v. 20, p. 366.
45. Raevsky O.A., Schaper K.-J., Artursson P., McFarland J.W. Ibid., 2001, v. 20, p. 402.
46. Goodford P.J. J. Med. Chem., 1985, v. 28, p. 849.
47. Boobbyer D.N.A., Goodford P.J., McWhinnie P.M., Wade R.C. Ibid., 1989, v. 32, p. 1083.
48. Wade R.C., Clark K., Goodford P.J. Ibid., 1993, v. 36, p. 140.
49. Wade R.C., Clark K., Goodford P.J. Ibid., 1993, v. 36, p. 148.
50. Goodford P.J. J. Chemomimetics, 1996, v. 10, p. 107.
51. Garon G., Rey S., Ermondi G., Crivori P., Guillard P., Carrupt P.-A., Testa B. In: Pharmacokinetic Optimization in Drug Research. Eds. B.Testa, van de H.Waterbeemd, G.Folkers, R.Guy. Weinheim: Wiley-VCH, 2001, p. 513.
52. Raevsky O.A., Skvortsov V. J. Comput.-Aided Molecular Design, 2002, v. 16, p.1.
53. 3D HYBOT, Registration by Russian State Patent Agency № 2004612207 of 27.09.2004
54. Raevsky O.A., Skvortsov V.S. SAR & QSAR in Environmental Research, 2005, v. 16, p. 287.